

# Apuntes sobre estadísticas



Preparado por:  
Armando Torres Pesaresi, M.S.  
Candidato al Grado de Doctor en Psicología (Psy.D.)

Editado por:  
Javier I. Toro Torres, Ph.D.  
Psicólogo Clínico

## Análisis de las hipótesis estadísticas

- En términos de análisis, luego de que se han establecido las hipótesis del estudio, se determinan las pruebas o test de hipótesis que probarán si lo que se planteó en las hipótesis fue lo que ocurrió en la realidad. El análisis estadístico que se hace con estas pruebas (por ej. **t** de student, **z**, **x<sup>2</sup>**, ect.) no son otra cosa que **procedimientos para tomar una decisión, bajo incertidumbre**, sobre la validez de la hipótesis nula usando la evidencia de los datos.

- **Error Alfa y Error Beta:** puesto a que trabajamos bajo incertidumbre es claro que cualquiera sea la decisión que tomemos siempre existe una probabilidad de cometer error. A fin de clarificar esto podemos presentar el siguiente esquema:

### Esquema del procedimiento

Decisión	Realidad sobre $H_0$	
	Cierta	Falsa
Rechazar $H_0$	<b>Error Tipo 1 (rechazarla cuando es cierta)</b>	Decisión correcta
No rechazar $H_0$	Decisión correcta	<b>Error Tipo II</b>

Como se puede ver en el esquema, con cada tipo de decisión que se tome hay asociado una posibilidad de cometer un error. Un procedimiento de este tipo sería óptimo cuando las probabilidades de cometer un error, cualquiera sea la decisión que se adopte, sean pequeñas. Lamentablemente, en la mayoría de las veces que se intenta probar las hipótesis, sólo es posible controlar una de ellas, con la circunstancia agravante de que estos errores son competitivos, es decir, cuando se disminuye mucho la probabilidad de uno aumenta la probabilidad del otro.

### Probabilidad del Error Tipo I- Error Alfa ( $\alpha$ )

- Puesto que, el interés generalmente es “rechazar  $H_0$ ” la probabilidad de error que se controla durante este procedimiento, es justamente el error asociado a esta decisión, es decir, **la probabilidad de rechazar  $H_0$  cuando es cierta**. La máxima probabilidad de error Tipo 1 se denota con  $\alpha$  y recibe el nombre de **nivel de significación de la prueba** y el debe ser prefijado de antemano (se hipotetiza y luego se confronta con la realidad).

### Probabilidad de Error Tipo II- Error Beta ( $\beta$ )

- La probabilidad de Error Tipo II se denota con  $\beta$  y es útil para encontrar la bondad del test que se mide en términos de la cantidad  $1 - \beta$  denominada **Poder de la Prueba**.
- El nivel de significación que se usa generalmente es  $\alpha = 0.05$  lo que corresponde a un 5% en términos de porcentaje.

## Ejemplo para comprender la aplicación del nivel de significación y poder de la prueba

- Si uno está probando el efecto curativo de un medicamento vs. un placebo, uno interesa validar su H1 tratando de establecer **diferencias bien marcadas**, que sea evidente el efecto del medicamento sobre el sujeto. Para ello, se establece un margen de error bien pequeño- que la posibilidad de equivocarme sea mínima. En el idioma estadístico se establece un nivel de significancia " $\alpha$ " de .05, o de .01, o de .001, entre otros.
- Por otro lado a veces lo que estas tratando de investigar, el **efecto dañino** o secundario de un tratamiento como su H1. Por lo que analizas la hipótesis de manera que, a la mera presencia de un efecto, o las **mínimas diferencia** se valida la hipótesis- que la posibilidad de equivocarme sea mayor (se escoge una prueba que sea "poderosa" en detectar estas mínimas diferencias). En el idioma estadístico se establece un nivel de significancia " $\alpha$ " de .10, o de .20, o de .25, entre otros.

### P-value

- Otro procedimiento general de una prueba de hipótesis **más usado en la actualidad** debido a la posibilidad de programas estadísticos (ej. SPSS), consiste en tomar la decisión a partir de la probabilidad del error Tipo 1 que brindan los "outputs" de tales programas, denominado **P-value** o simplemente P. Este procedimiento lo podemos resumir de la siguiente manera:
    - ✓ Si  $P \leq \alpha$  entonces la decisión es **Rechazar Ho**
    - ✓ Si  $P > \alpha$  la decisión es **No hay evidencia suficiente para rechazar Ho (se acepta/retiene)**.
  - **Ejemplo 1:** si estableció su hipótesis con un nivel de significancia de .01 ( $\alpha = .01$ ), y al realizar los análisis estadísticos le resulta un p-value de .02 ( $p = .02$ )  
 **$P > \alpha$ ; .02 > .01;** por lo tanto- no hay suficiente evidencia para rechazar Ho
  - **Ejemplo 2:** si estableció su hipótesis con un nivel de significancia de .01 ( $\alpha = .01$ ), y al realizar los análisis estadísticos le resulta un p-value de .001 ( $p = .001$ )  
 **$P < \alpha$ ; .001 < .01;** por lo tanto- se rechaza la Ho
- 

## Medidas de Tendencia Central o posición

- Las medidas de posición nos facilitan información sobre la serie de datos que estamos analizando. Estas medidas permiten conocer diversas características de esta serie de datos. Informan sobre los valores medios de la serie de datos.
  - **La media aritmética:** es una medida de resumen que considera todos los valores de una distribución bajo estudio. Lo que hace es buscar un punto medio considerando **el valor** de cada uno de los datos. Es una de las medidas que más se utiliza para inferir información respecto a unos datos. Esta medida tiene 2 limitaciones:
    1. **Se afecta por valores extremos:** esto es, se afecta su interpretación (el referirse al punto medio) cuando existen valores anormales a los otros.

Ejemplo: un estudio con una prueba que mide motivación en empleados públicos, la mayoría ofrece puntuaciones bajas (poca motivación) y siempre alguno (muy a gusto con su trabajo) presenta una puntuación extrema mayor.

2. **No puede obtenerse cuando los datos están en una distribución con clases abiertas:** esto es, si esta escala de intervalo, se da una situación en la que se dividen los grupos: 5 o menos, entre 6 y 10, entre 11 y 15, 16 o más. Aquí se presenta que entre 5 o menos puede que el menos sea "0", pero en 16 o más puede que sea 100, 299, ect.

- **La mediana:** es el punto que en que la distribución se divide en dos partes iguales, o sea, el 50% de los datos. Esta no se afecta por valores extremos. No considera el peso del valor, sino el orden de los datos de la distribución para ubicarla. **Por lo tanto, en aquellas situaciones donde se observa considerable asimetría, la mediana es la mejor indicación de la tendencia central que la media.** Se localiza a través de posición:

**Ejemplo:** Para los datos 3,5,7,7,8,9

**El promedio:**  $(3+5+7+7+8+9) / 6 = 6.5$

**La mediana:**  $\# \text{ casos} + 1 / 2 = 6 \text{ casos} + 1 / 2 = 7/2 = 3.5$

**Posiciones:**

1	2	3	4	5	6
3	5	7	7	8	9

Para determinar la posición 3.5 hay que sacar el promedio del dato en la posición 3 y el dato en la posición 4. En este caso, como es el mismo dato, la mediana es 7.

- **Moda:** basta con reconocer o recordar que es el dato que más se repite en una distribución. Es un análisis rápido de tendencia central pero es el menos científico. Se utiliza en aquellos casos donde se requiere de improvisa una medida central. Cuando se observa que varios datos se repiten en la misma cantidad se habla de que existe más de una moda (multiple mode). En el ejemplo anterior la moda era el 7.

### En resumen:

<b>Media aritmética (Media)</b>	Medida que considera todos los valores de una distribución bajo estudio. Busca el punto medio considerando el valor de cada uno de los datos. Una de las medidas más utilizadas para inferir información respecto a unos datos. Se afecta por datos extremos y no se puede usar en distribuciones con clases abiertas.
<b>Mediana</b>	Es el punto en que la distribución se divide en dos partes iguales (50% de los datos). Se considera el orden de la distribución para ubicarla. En situaciones donde se observa asimetría, la mediana es una mejor indicación central de la medida.
<b>Moda</b>	Es el dato que más se repite en una distribución. Si se observa que se repiten varios datos en la misma cantidad se habla de que existe más de una moda.

---

### Medida de posición no centrales (Medidas de Dispersión o Variabilidad)

- Permiten conocer otros puntos característicos de la distribución que no son los valores centrales. Entre otros indicadores se suelen utilizar una serie de valores que dividen la muestra en tramos iguales:
  1. **Cuartiles:** son tres valores que distribuyen la serie de datos ordenadas de forma creciente o decreciente, en cuatro tramos iguales, en los que cada uno de ellos concentra el 25% de los resultados.  
0%....**(1cuartila)**.... **25%**.....**50%**.....**75%**.....100%
  2. **Deciles:** son 9 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en diez tramos iguales, en los que cada uno de ellos concentra el 10% de los resultados.  
0%...**10%**....**20%**....**30%** ....**40%**....**50%**....**60%**....**70%**....**80%**....**90%**...100%
  3. **Percentiles:** son 99 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cien tramos iguales, en los que cada uno de ellos concentra el 1% de los resultados.
  4. **El Rango Percentil (Percentil Ranks):** compara un resultado del estudiante con una muestra nacional de estudiantes en el mismo grado. El rango percentil indica el porcentaje de estudiantes en la nación que lograron una puntuación igual o más baja que su hijo. Por ejemplo, Ashley con un rango percentil de 38% logró una puntuación mejor que el 38 % de los estudiantes que se examinaron de ese grado. El resto de los estudiantes, 62% anotaron igual o superaron a su hija.

**REPASAR PAG 10 y 11**

---

## Pruebas estadísticas

- En el ambiente de la investigación usted hace comparaciones; según lo que está comparando se escogerá una prueba estadística que se ajuste **al tamaño de la muestra, la escala** en que están los datos y **tipo de comparación** o contraste que se realiza. Utilizando términos del idioma estadístico, usted puede realizar comparaciones entre una **muestra** (por ejemplo: su moral) con el **universo** (la moral general, **Muestra vs Universo**); puede realizar comparaciones **entre una muestra** (ej: usted como profesional) **vs. otra** (sus pares de profesión) conocida como **Entregrupos**; puede realizar comparaciones **en un mismo grupo en dos momentos distintos** (por ejemplo: tus conductas adolescentes vs. tus conductas de adulto; **Intragrupo o Muestras Correlacionadas**).

### Contrastes entre la media de una muestra con la media del universo

- Curva Normal:** la distribución normal es una de las distribuciones más apropiadas; siempre que la muestra está constituida por **30 o más** participantes. Toma una forma de campana perfectamente simétrica.

Para verificar rápidamente si su Z obtenida es significativa: **Z crítica** si la prueba es unilateral o bilateral. **Interpretación:** si su Z obtenida es mayor que la señalada de significancia, habrá diferencias entre la muestra y el universo.

Nivel de significancia	Prueba unilateral	Prueba bilateral
.05	Z= 1.645	Z= 1.96
.01	Z= 2.33	Z= 2.58
.001	Z= 3.08	Z= 3.30

---

### Modelos T para contrastes entre la media de una muestra con la media del universo

- A diferencia de la curva normal, los modelos T se utilizan cuando se desconoce la desviación estándar del universo y cuando las **muestras son de 30 casos o menos**.
- Grados de libertad:** se resume para los efectos de la siguiente manera= **n-1**  
Usted restará al tamaño de su muestra un caso, y ubicará ese valor en la tabla de valores críticos, cruzándolo con el nivel de significancia escogido.
  - Así que si **le preguntan los grados de libertad para una comparación como esta:** n-1. Si n=10, grados de libertad =9.
  - Interpretación:** si el valor T obtenido **es menor** (el número obtenido) **que el valor T crítico** (según el nivel de significancia), **no habrá diferencias que sean significativas**, esto, es se pueden explicar por causalidad o error en el estudio.
  - Si el valor T obtenido **es mayor** que el valor **T crítico** (según el nivel de significancia), **habrá diferencias significativas** entre la muestra y el universo que se pueden relacionar al efecto de las variables estudiadas.

### CHI cuadrado para contrastes de la varianza de una muestra con la varianza del universo

- Esto es similar al contraste de una media de una muestra y la media del universo. En ambos casos se formulan hipótesis, se selecciona el nivel de significancia, se sustituye en la fórmula y se examina el resultado contra los valores críticos.
- Se obtendrá  $X^2$  que se comparará con la  $X^2$  crítica para analizar su significancia. En el caso de utilizar un  $X^2$  para comparar varianzas.
- Si la hipótesis planteada en términos de ">" (se espera que la varianza de la muestra **sea mayor** que la del universo), la  **$X^2$  obtenida será significativa si es igual o mayor que la  $X^2$  crítica.**
- Si la hipótesis es planteada en términos de "<" (se espera que la varianza de la muestra **sea menor** que la del universo), **la  $X^2$  obtenida será significativa si es igual o menor que la  $X^2$  crítica.**
- Pero, si la hipótesis es planteada en *términos de  $\neq$ , corresponde a una prueba bilateral*. Si su nivel de significancia para su hipótesis es de .05, ( $0.5/2 = .025$ ) los valores críticos para  $X^2$  deben ubicarse para ambas direcciones. El valor de  $X^2$  obtenida será significativo si es mayor que la  $X^2$  crítica para un  **$\alpha = .025$  o menor de  $\alpha = .975$ .**

---

### Modelos T para contrastes de medias para muestras correlacionadas

- Un diseño de muestras correlacionadas es aquel donde se mide la variable que nos interesa antes y después de una situación experimental o tiempos distintos.
- Se conoce también como diseños **intra-grupos**.
- El procedimiento a seguir en estos diseños: se formula la pregunta de interés y luego se establece la hipótesis de nulidad.
- Cuando se aplica el modelo T al diseño intra-grupo, la hipótesis nula se representa de la siguiente manera: la diferencia entre las medias es igual a 0 y la suma de las diferencias para cada uno de los sujetos, la cual se dividirá entre el tamaño de la muestra.
- La  $H_0$  indica que la diferencia entre las medias antes y después es igual a 0. Deben entender que en un estudio correlacional, se realiza una medición de una variable de interés para el investigador en por lo menos 2 tiempos, pero su interés es manipular los sujetos de manera tal que al medir nuevamente esas variables el resultado sea distinto al primero.
- De forma semejante a lo discutido en el contraste de la muestra vs su universo, aquí se desea encontrar diferencias significativas. Pero, se parte de que las diferencias sean distintas a 0, mayores que 0 o menores que 0:
  - Si esperamos que la primera medición (antes) sea mayor que la segunda (después), la  $H_1: D > 0$ .
  - Si esperamos que la primera medición (antes) sea menor que la segunda (después), la  $H_1$  se representaría de la siguiente manera:  $H_1: D < 0$ .
  - Si no tenemos suficiente evidencia para conocer qué dirección tomarán las diferencias, expresamos la hipótesis alterna de modo bilateral:  $H_1: D \neq 0$
  - Luego de tabular estos datos aplicará la siguiente fórmula:  $t = D / sD$

---

## CHI cuadrado (X2) para pruebas de proporciones: diseños antes- después (X2)

- En algunos estudios es de interés examinar cómo se modifican la proporción de ciertas variables en los mismos sujetos una vez expuestos a varias intervenciones.
- Para nuestro interés nos limitaremos a 2 intervenciones.
- La prueba consiste en determinar **si la proporción que responde de cierta manera antes se modifica después de la intervención**. De esta manera, la hipótesis se puede representar de la siguiente manera:  $H_0: P_1 = P_2$  /  $H_1: P_1 \neq P_2$

---

## Contraste para muestras correlacionadas cuando la escala es ordinal

- Existen situaciones donde las características bajo estudio aparecen tratadas en términos de escalas ordinales: donde los números o clases reflejan el **orden o secuencia de las particularidades involucradas**.
- Bajo estas circunstancias existen pruebas estadísticas distintas a las antes vistas.
- Por ejemplo, una actitud que favorece más o menos X situación, medir si una conducta es más o menos agresiva, mayor o menor prejuicio, mayor o menor cantidad de síntomas de X diagnóstico, ext.
  - **Prueba de signos y distribución binomial**: una de las aplicaciones para escala ordinal cuando el diseño es **antes y después** es esta prueba, la cual no utiliza valores numéricos reales sino los signos "+" y "-" como indicadores de aumento o reducción, sin considerar la magnitud de las diferencias.
  - Aplicación de esta herramienta estadística: que la muestra sea pequeña (menos de 30 casos) se asignará el signo positivo cuando al comparar antes y después haya un aumento en después se asignará el signo negativo cuando al comparar antes y después haya una disminución en después.
  - La probabilidad obtenida se contrasta con el nivel de significancia establecido. Si es igual o menor que el nivel de alfa establecido, se rechaza la hipótesis nula. Si es mayor, se acepta la nula.
- ✓ **Prueba T de Wilcoxon**: a diferencia de la prueba de signos (prueba anteriormente discutida), la T de Wilcoxon considera la magnitud del cambio observado en la escala ordinal entre el momento antes y el momento después (es más potente).
- ✓ Esto significa que en un momento dado podríamos aceptar la  $H_0$  aplicando la prueba de signos y rechazarla utilizando TW. Tanto así que, si la prueba de signos arroja resultados significativos, TW también conducirá a la misma conclusión.
- ✓ Pero lo inverso que si TW arroja resultados significativos, la prueba de signos no necesariamente los arrojará.
- ✓ Pasos para aplicar esta prueba:
  1. Las muestras deben de tener un **tamaño entre 5 y 50 casos**.



2. Se requiere identificar la diferencia entre antes y después junto a la dirección del cambio.
  3. A cada diferencia se le va a asignar un rango, el cual representa una posición partiendo desde la menor hasta la mayor. El rango se identifica con el signo que tiene del cambio.
  4. Se suman por separado los rangos positivos y los negativos.
  5. La suma menor absoluta será la TW, la cual se contrastará con el valor crítico que puede ubicar en la Tabla G/ apéndice I (libros de Sánchez Viera).
  6. Si TW obtenida es igual o menor que la TW crítica, se rechaza la hipótesis nula. Si TW obtenida es mayor que la TW crítica, se acepta la hipótesis nula.
- 

### **Contraste entre la media de 2 muestras**

- **Utilizando curva normal (z):**
  - Además de contraste entre la media de una muestra y la media del universo, la curva normal puede emplearse en pruebas donde se comparan los datos obtenidos de dos grupos distintos, o independientes.
  - Recuerde que esta se utiliza cuando los grupos a compararse incluyen **30 o más casos** cada uno. Para utilizar esta herramienta estadística, **se compararán las medias de cada uno de los grupos;** a base de esta consideración, la hipótesis estadística podrá ser redactada cómo se presentan a continuación: Hipótesis nula= promedio 1 = promedio 2 / hipótesis alterna= promedio 1 ≠ a promedio 2.
- 

### **Escala nominal: CHI 2 para pruebas de independencia/ no paramétricas**

- Cuando los datos están presentados en escalas nominales, no requiere que la información parta de parámetros.
- En las Ciencias Sociales la aplicación de la estadística como Dios manda es un poco cuesta arriba, ya que muchas de las preguntas de investigación pueden contestarse solamente considerando muestras en su estado natural, difíciles de identificar o clasificar, por lo que su distribución en la población se desconoce.
- X<sup>2</sup> es la técnica más utilizada cuando los datos no son paramétricos, y la observación es considerada en categorías.
- Por lo tanto, utilizaremos los conceptos enseñados y practicados anteriormente, pero modificados para este tipo de comparación.
- En este caso, ya que no existe un parámetro para comparar, se contrasta la cantidad de veces que se observa un dato con la cantidad de veces que es esperado este dato.
- Lo observado lo obtenemos; lo esperado es un cálculo que se realiza a partir del tamaño de la muestra.
- La sumatoria de la diferencia cuadrada entre lo observado (o) y lo esperado (e)/ entre lo esperado (e).

- Las hipótesis estadísticas se presentan de la siguiente manera:  $H_0: \mu = e$  /  $H_1: \mu \neq e$
- 

### **Prueba de la mediana, para K muestras independientes**

- Esta prueba también se recomienda para escalas ordinales donde el análisis de varianza puede verse afectado por violaciones al supuesto de normalidad.
  - Permite contrastes entre K número de grupos tomando como punto de referencia los valores críticos de  $\chi^2$ .
  - En la pruebas de las medianas las hipótesis se plantean de la misma forma que en  $\chi^2$ .
  - Si los grupos provienen del mismo universo, la mediana en cada grupo no debe diferir significativamente de la mediana obtenida para todos en conjunto.
  - Siendo la mediana una medida que divide la distribución en dos partes iguales, si la hipótesis nula es cierta esperamos la mitad de los sujetos sobre dicho indicador y la otra mitad por debajo.
  - **$H_0$ : establece que los grupos provienen de un mismo universo con una mediana común a todos.**
  - **$H_1$ : establece que los grupos no provienen del mismo universo, o sea los sujetos de los grupos que se están considerando difieren en relación a la variable que se está considerando.**
- 

### **Prueba H de Kruskal-Wallis:**

- Se utiliza para los mismos propósitos que la mediana, es más potente que la prueba de la mediana, o sea es más probable rechazar la  $H_0$  nula cuando es falsa con esta que con la de la mediana.
  - Se usa en escalas ordinales.
  - Al igual que practicamos en la TW se organiza la información catalogándola en rangos a partir de valor menos o mayor, considerando todas las muestras como una.
  - **La aplicación de esta nos permite determinar si las muestras proceden de universos similares o no.**
  - Si son similares, la media de los rangos será aproximadamente la misma en cada grupo.
  - Si proceden de universos distintos, el índice (H) será significativo.
  - Utiliza como tabla de referencia para determinar la significancia  $\chi^2$ .
  - Las Hipótesis estadísticas se describen de la siguiente manera:  $H_0: H = 0$  /  $H_1: H \neq 0$
- 

### **Correlación**

- El grado de asociación entre las variables se mide en función de los llamados coeficientes o índices de correlación.
- En algunos casos dichos índices fluctúan entre 0 y +1.

- En otros casos van desde -1 a +1 (aquí se puede hablar de dirección de la asociación, sea inversa o directa).
- La fluctuación entre -1 y +1 se presenta solamente en investigaciones donde se comparan escalas ordinales, de intervalo o razón.
- Si las escalas son nominales no hay una gradación numérica, por lo que el resultado que se obtiene no puede interpretarse como indicando relación directa o inversa; siempre son positivos.
  - **Escalas nominales:** el resultado solo puede señalar la intensidad de la relación, pero no la dirección.
  - Comparaciones para tablas 2 x 2.
  - **Phi ( $\phi$ ):** fluctúa entre 0 y 1
  - Según el coeficiente se acerca a 0, más baja será la relación; según se acerca a 1, más intensa será la relación.
  - **Lambda:** fluctúa entre 0 y 1
  - A más cerca del 1, mayor intensidad de la relación, y a más cerca a 0 menor intensidad de la relación.
  - **Interpretación:** el % de reducción en error al predecir la ocurrencia de una cosa partiendo de la variable dependiente establecida.
  - A más bajo el Lambda, mayor será el error de predicción.

---

#### **Coeficiente para tablas mayores de 2 x 2:**

- Significa que por lo menos una de las variables que tiene 2 o más alternativas.
- Aquí se utiliza Lambda y la V de Cramer (si una o ambas escalas son nominales).
- Ejemplo: Resultado V.59/ Lambda .439
- Interpretación: (ver tabla página 20) el resultado nos sugiere una relación moderada alta entre la afiliación política de la facultad y la actitud hacia la presencia de la marina en Vieques puede notarse que los facultativos del PIP y del PPD muestran una actitud más favorable a la frase, mientras que los del PNP y los afiliados al ningún partido tienden a desfavorecer este “slogan”.
- El índice obtenido nos indica que el % de reducción en error es la predicción de la actitud conociendo la afiliación política de la facultad es 43.9
- **Escalas ordinales:** es posible observar relaciones directas o inversas, siempre y cuando la relación sea lineal.
- Coeficiente más utilizado para estas: Spearman ( $r_s$ ). Correlaciona rangos o posiciones, no la magnitud de las variables.
- Si el coeficiente obtenido es positivo se interpretará la relación como una directa (a mayor ocurrencia de x evento, mayor probabilidad de que ocurra y evento).
- Si es negativo se interpretará como inversa (a mayor ocurrencia de x evento, menor ocurrencia de y evento).
- Otro coeficiente en esta escala: Gamma ( $g$ .) fluctúa entre 0 y 1 si la relación es directa y entre 0 y -1 si la relación es inversa.

- Puede aplicarse comparando los resultados de las escalas directamente y también cuando las variables se representan a través de categorías ordinales (rs no puede usarse aquí).
  - Ejemplo: un investigador estaba interesado en determinar si había relación entre el alcohol y satisfacción marital. Utilizó dos instrumentos de escala ordinal para operacionalizar las variables y obtuvo lo que aparece a continuación (Tabla pág. 20), donde a más lato índice, mayor consumo de alcohol y mayor insatisfacción marital.
  - Resultado  $r_s = 9.85$ / Interpretación: el índice obtenido sugiere una relación directa sumamente fuerte entre el nivel de consumo de alcohol y la puntuación en la escala de satisfacción marital. Puede encontrarse una tendencia de a mayor consumo de alcohol, mayor insatisfacción marital. Gamma= .96
  - Interpretación: sugiere una intensa relación directa entre el consumir alcohol y satisfacción marital (se interpreta de forma similar).
  - **Escala de intervalos/ razones:** el índice de correlación que se aplica se denomina correlación producto momento de Pearson ( $r$ ). Esta solo puede utilizarse cuando la relación es lineal. ***Directa entre 0 y 1; inversa entre 0 y -1.***
  - Ejemplo: un psicólogo industrial tomó una muestra de empleados y tomó el número de días que se ausentaron durante 6 meses, así como el nivel de involucración en el trabajo. Obtuvo lo siguiente: resultado:  $r = -.91$ / Interpretación: la correlación obtenida sugiere que entre las ausencias y la involucración al trabajo hay una relación inversa fuerte, estos es, a mayor involucración, menos ausencias.
-

Herramienta	Utilización e Interpretación
<b>Alcance o amplitud</b>	Índice más simple. Es un dato para comparar cuán amplio fue el recorrido de una distribución vs la otra. Es la diferencia entre el dato de mayor y de menor valor numérico.
<b>Media aritmética</b> (sumas todos valores y los divides entre la cantidad de personas)	Medida que considera todos los valores de una distribución bajo estudio. Busca el <b>punto medio</b> considerando el valor de cada uno de los datos. Una de las medidas más utilizadas para inferir información respecto a unos datos. Se afecta por datos extremos y no se puede usar en distribuciones con clases abiertas.
<b>Mediana (Md)</b> (tomas los valores y le sumas 1 y divides entre 2, luego localizas la posición de 50%)	Es el punto en que la distribución se divide en dos partes iguales ( <b>50% de los datos</b> ). Se considera el orden de la distribución para ubicarla. En situaciones donde se observa simetría la mediana es una mejor indicación central que la media.
<b>Moda (Mo)</b>	Es el dato que más se repite en una distribución. Si se observa que se repiten varios datos en la misma cantidad se habla de que existe más de una moda.
<b>Desviación estándar</b> <b>S, <math>\sigma</math></b>	Medida de referencia para saber cuánto difieren los datos y cuánto se asemejan. Es un tipo de agrupación que considera el nivel de separación de los datos. El valor de la desviación estándar se sumará y restará al valor obtenido de la media aritmética para así comprar la dispersión de los datos a comparar. La desviación estándar para la muestra es <b>S</b> y para el universo es <b><math>\sigma</math></b>
Varianza ( $s^2 / \sigma^2$ )	Representa la homogeneidad y heterogeneidad de los datos. En términos de la varianza, los datos mas homogéneos tendrán una varianza más pequeña y los heterogéneos una más grande. La varianza para la muestra es $S^2$ y para el universo $\sigma^2$ .